Title: METHOD FOR DETERMINING AND ELIMINATING THE DRIVERS OF NON-VALUE ADDED COST DUE TO PRODUCT COMPLEXITY AND PROCESS PARAMETERS

CROSS REFERENCE TO RELATED APPLICATION

[0001]   This Application claims the priority of U.S. Provisional Patent Application Serial No. 60/479,333, Filed: June 18, 2003.

BACKGROUND

[0002] Organizations that provide services or manufactured products constantly face a fundamental precept of business, that is to reduce the costs of providing a highly desirable product or service in order to maintain a Return on Investment attractive to the owners.

[0003]   These forces have caused companies to reduce internal costs using continuous improvement processes (e.g. Six Sigma, Lean, etc) to eliminate non-value added costs associated with quality and long lead times. Non-value added costs are defined as those costs that do not add a form, feature or function that the customer values. Examples include the cost of obsolescence, scrap and rework, excess capital investment, costs associated with transportation, stockrooms, warehousing, scheduling and supervision, and low direct labor productivity.

[0004]   Businesses have also expanded the breadth of their product offering to grow revenue and respond to market demand.   However, the costs associated with increasing complexity of products and services offered results in non-

1

value added costs that have hitherto been difficult or impossible to estimate relative to other non-value added costs. Some of this complexity has resulted from the historical addition of products and services with no thought of standardizing subcomponents. Some of the cost of complexity is due to the propensity to add new products without a corresponding effort to eliminate products. The lack of symmetrical effort is in part due to a lack of knowledge of the cost of complexity. This has inhibited businesses from eliminating products that add cost but do not earn an adequate return on investment. This invention provides methods to make a quantitative determination of the costs of complexity, versus those due to quality, long setup time, downtime, etc. This will make it possible for business entities to make rational investments in complexity reduction along with other process improvement initiatives. The process of determining the cost of complexity applies to manufactured products as well as service items. Because manufactured products are tangible, such will be used for purposes of initial exposition and then generalized to service processes.

[0005] The cost of product complexity arises in two dimensions. Business entities that have designed products or services often, over time create unique components, part numbers or subassemblies for each end item offered to customers. As the number of end items grows, so do the

number of unique components and the associated complexity costs. With suitable redesign, several offerings or end items can use common components or subassemblies, reducing the cost of internal complexity which is transparent to the customer. This is referred to as Non-value Add complexity. For example, the first named inventor herein was the CEO of International Power Machines. Over a period of five years, eight separate mechanical and electrical designs had grown to provide uninterruptible power supplies with ratings of 10, 20, 25, 37.5, 50, 60, 80, and 100KW. By suitable redesign, more than 90% of all mechanical and electrical components and subassemblies were made common. This is an example of non-value added complexity, because the customer derived no benefit from the internal differences. In the present invention equations have been derived which compute the excess inventories, learning curve, and related costs that were eliminated by this process, resulting in an increase in gross profit margin from 15% to 37%. Each standardized rating adds about 20 unique part numbers. Standardization reduces the number of incoming supplies from fewer vendors, driving larger volumes through fewer items also resulting in lower cost. This process of standardization is depicted in FIGURE 1.

[0006] In principle, it is possible to design a platform which will pre-standardize the components of the rollout of a new family of products. Standardization has the merit of

making breakthrough cost reductions without eliminating any customer facing products or services. The process of producing any product or service can be analyzed as a process which links activities where work is done. Some of this work adds value as perceived by the customer. Many activities, such as moving material, operating stock rooms, warehouses, performing rework, re-learning work, etc., add no value. Many of these costs are due to complexity of the product.

[0007] A second cost of complexity dimension is value optimization and customer facing complexity. Standardization that simply eliminates non-value added complexity can be executed once the costs are known. However, products whose features and functions are desired by the customer, i.e., value added complexity, may not earn their cost of capital and should be considered for removal, out-sourcing, re-pricing, or process improvements which will eliminate the costs of complexity. But the true cost of carrying the complexity of low volume products may not be disclosed by traditional accounting. The true cost may be disclosed by activity based costing (e.g. Kaplan), but may logically lead to the elimination of products that do not earn their cost of capital. This invention provides prioritized process improvement methods for eliminating the cost of complexity and retaining a potentially valuable source of business owner return. Analysis in accordance

with the invention will also determine if a product or service cannot be suitably improved, within the market price, leading to its elimination to optimize business owner value. By ordering the Stock Keeping Units (SKU) by decreasing Return On Invested Capital (ROIC), a curve similar to FIGURE 2 has resulted.

[0008] A purpose of the invention is to enable a business entity to estimate the relative benefits of complexity reduction such that rational investments can be made versus other priorities.

[0009] Essentially any business process can be analyzed as a series of steps or activities which, in the aggregate, add significantly to the cost of producing the product or providing the service. Certain value added activities in business processes are, of course, essential to providing the product or service desired by the customer and are characterized as value added steps. These steps add a form, feature or function to the offering or product for which the customer would be willing to pay. However, certain non-valued added activities may contribute cost into the overall process as a result of process deficiencies or result from the complexity of the product or service itself. Process deficiencies such as the defects due to variation in quality will create non-value added activities of rework, or, if scrapped, will cause additional replacement products to be produced. In service environments, correcting an invoice,

implementing an engineering change notice, etc. are examples of non-value added activities. In both cases, the non-value added activities act to slow down the velocity of the product or service. According to Little's Law, for a given level of demand, a slowdown in cycle time will increase the amount of Work In Process.

[0010] Other process deficiencies are directly responsible for slowing the velocity of the process. For example, if an activity has a long setup time to changeover from one customer requirement to another, it will cause the operator to "lock on" to customer requirements to minimize the wasted time in setup. However, this approach means that some customer requirement must wait a long time in queue, resulting again in large amounts of Work in Process, and creating major congestion and delay. This will cause excessive investment in Work in Process which creates non-value added costs of stockrooms, exacerbating the quality problems and, lost revenue. In service environments, this problem can often be mitigated by hiring additional labor to meet customer demands. Other time related process deficiencies include downtime or absenteeism, long process in time, poor work organization, queue time due to variation in demand, service, and supply rates, etc. Methodologies such as the Toyota Production System, Kanban, Lean, and Lean Six Sigma have been used to

accelerate process velocity and eliminate non-value added cost and excessive work in process.

[0011] The two process deficiencies related to quality and speed have the result of increasing the number of value added activities in a process and effectively lengthening the chain from input to output. Experience shows that the majority of activities in unimproved processes are non-value added.

[0012] There is a third independent source of non-value add cost, slow velocity and large Work In Process (WIP) that is due to the complexity of the service or product offering itself. This source has been qualitatively understood and is the subject of many articles. What has not been understood is how to quantitatively measure the amount of non-value added cost due to product complexity versus the non-value added cost due to quality and speed deficiencies. While improvements in all three areas are important, their relative importance must be quantitatively understood to make rational investments in process improvement. The present invention provides an unambiguous method that allows this determination. FIGURE 3 illustrates the relative power of Lean, Six Sigma and Complexity Reduction in reducing the amount of WIP.

SUMMARY OF THE INVENTION

[0013] The present invention provides a method for measuring the effect of product or service complexity on

non-value added cost in processes for producing products and services, relative to the non-value added costs created by process deficiencies related to process speed or quality. The complexity of a product or service is the result of (a) too many products or services that are sold to customers which overlap or do not earn an adequate Return On Invested Capital and should be reduced through rationalization and (b) too many different modules or subassemblies which are used to create products or services that are sold to customers which can be reduced through standardization. In many business enterprises, this unnecessary complexity is the result of products or services being introduced over time with no attempt made to either rationalize or standardize the products.

[0014]   In accordance with the present invention a method for measuring and reducing the cost of complexity in a product manufacturing process is provided which utilizes a continuous improvement process formula wherein Work In Process and finished goods in inventory are reduced by taking into consideration such factors as manufacturing setup time, the percentage of finished goods which are defective, the number of different parts, processing time per unit of product, workstation downtime, rework, variation in demand and supply rates and other variables which may be taken into account. Less well known is the fact that the cost of non-manufacturing service processes suffer from the

same causal factors of setup time, variation in demand, etc., and hence are susceptible to the same form of analysis.

[0015] Other facets, advantages and superior features of the present invention will be appreciated by those skilled in the art upon reading the further description which follows in conjunction with the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] FIGURE 1 is a diagram indicating the improvement in gross margin of a product line before and after standardization;

[0017] FIGURE 2 is a diagram indicating the economic profit versus the number of products and showing the value optimization zone;

[0018] FIGURE 3 is a diagram indicating the relationship between reduction of setup time or complexity as a function of the ratio of reduced Work In Process to initial Work In Process;

[0019] FIGURE 4 is a process flow diagram illustrating the relationship and presence of value added and non value added steps;

[0020] FIGURE 5 is a diagram indicating the cost of goods sold versus adjusted Work In Process turns;

[0021] FIGURE 6 is an illustration of a typical value stream map;

[0022]   FIGURE 7 is a diagram showing a typical workflow from raw material deliveries to deliveries of finished goods to a customer;

[0023]   FIGURE 8 is a diagram indicating a synchronized product schedule in accordance with the invention;

[0024]   FIGURE 9 is a diagram illustrating raw material inventory in front of a first workstation with synchronized production;

[0025]   FIGURE 10 is a diagram indicating Work In Process inventory between first and second workstations with synchronized production;

[0026]   FIGURE 11 is a diagram indicating finished goods in inventory with synchronized production;

[0027]   FIGURE 12 is a diagram indicating aggregate inventory levels with synchronized production;

[0028]   FIGURE 13 is a diagram indicting a setup-on-batch-arrival production schedule;

[0029]   FIGURE 14 is a diagram indicating raw material inventory in front of the first workstation with a setup-on-batch-arrival schedule;

[0030]   FIGURE 15 is a diagram indicating Work In Process inventory between workstations with a set-up-on-batch-arrival schedule;

[0031]   FIGURE 16 is a diagram indicating finished goods inventory with a setup-on-batch-arrival schedule;

[0032]   FIGURE 17 is a diagram indicating aggregate inventory levels with a setup-on-batch-arrival schedule;

[0033]   FIGURE 18 is a table indicating a moderately asynchronized production schedule;

[0034]   FIGURE 19 is a diagram showing raw material inventory in front of a first workstation for a moderately asynchronized production schedule;

[0035]   FIGURE 20 is a diagram indicating Work In Process inventory between first and second workstations for a moderately asynchronized production schedule;

[0036]   FIGURE 21 is a diagram indicating finished goods inventory for a moderately asynchronized production schedule;

[0037]   FIGURE 22 is a diagram showing aggregate inventory levels for a moderately asynchronized production schedule;

[0038]   FIGURE 23 is a table showing a fully asynchronized production schedule;

[0039]   FIGURE 24 is a diagram showing aggregate inventory levels for a fully asynchronized production schedule;

[0040]   FIGURE 25 is a table showing a comparison of aggregate inventory levels for various schedules;

[0041]   FIGURE 26 is a table showing the total average Work In Process for a complex facility;

[0042]   FIGURE 27 is a continuation of the table of FIGURE 26;

[0043]   FIGURE 28 is a diagram showing the relationship between Fast Movers and Slow Movers as a function of Work In Process units;

[0044]   FIGURE 29 is a diagram indicating average backlog with demand variation;

[0045]   FIGURE 30 is a diagram indicating average backlog with setup/processing variations; and

[0046]   FIGURE 31 is a diagram indicating how lead time is lengthened by the learning curve when changing tasks.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

[0047]   Substantially all business processes can be represented as a series of steps or activities which are linked to provide a desired product or service. Referring to FIGURE 4, for example, the manufacture of a product can be represented by a source of raw or partially furnished material which flows through a value stream until it emerges as a product ready for market. Workstations (1), (2), (3), (4) and (5) are interposed the source (supplier) 10 and the product ready for market 12. Value added steps are carried out at stations (1), (2), (3), (4) and (5), as indicated by the symbols "VA" at each station.

[0048]   The ideal flow through the process of FIGURE 4 is represented by the dashed lines. However, certain non-value added steps 14, 16, 18 and 20 may be interposed the value added steps at workstations (1), (2), (3), (4) and (5) as indicated in the diagram. These non-value add steps and

their associated waste are in fact required to deliver goods and services at required completion rates despite the viscosity of the excess work in process. For example, material for work in process may require steps, such as placement of the material in stock rooms, movement to another workstation and/or testing. If a quality problem erupts, the amount of scrap and rework and, hence, the attendant non-value added activities will be directly proportional to the amount of Work In Process. In fact, the cost of non-value added steps is normally a major part of non-value add costs such as manufacturing overhead cost, and quality cost, which are typically two to three times the value added cost. A Complexity Value Stream Map (described below) segments each activity as value add, non-value add, or some mixture of both, and captures cost and process data. This mapping process often discloses that the number of non-value add activities in an unimproved process is significantly larger than the number of value add activities. The common metric that couples quality, speed and complexity and allows for the measurement of their relative impact on non-value add cost is their impact on the amount of Work In Process that they create. The relative value add reduction of process improvement versus complexity reduction can be determined analytically based on their relative reductions in Work In Process. It can be argued that a product that moved directly from value add to value

add activity would have no non-value add cost. The first instance of this occurrence was approximated by the Highland Park, Michigan Model T automobile factory. One of the principal causes of this low cost production was the boundary condition that only one product would be produced, eliminating all variety and, hence, complexity, setup time, etc.

[0049]   Empirically, a Tier 2 supplier of automotive coupled hoses was able to improve inventory turns (measured as Cost of Goods Sold (COGs/WIP) from 23 to 66 using a combination of process improvements and complexity reduction (internal part numbers were reduced by 35%). The result was that manufacturing overhead cost fell by 21%, or $6.0m out of a total of $28.6m. The relationship between WIP and cost is a fluctuating function as shown in FIGURE 5.

[0050]   It should be noted that Complexity Value Stream Mapping estimates the beginning and ending non-value add cost. The curve between is subject to all the exigencies of the improvement process and a single analytic curve is thus not to be expected.

[0051]   The increasing slope of cost reduction is due to the quantized nature of many costs such as off site warehouses, plant consolidations, outsourcing of non-core competencies, etc. which can only take place when the flow of work is essentially from one value add activity to another. This only occurs as cycle efficiencies (discussed

below) exceed 20% which relates to reduction of Work In Process of 60-90%.

[0052]   It also explains why a company, such as Toyota Corporation, strives for WIP turns and cycle times that are faster than required by the market by the application of Lean, Six Sigma tools.   To make the analysis revenue neutral, divide the WIP turns at any given moment by the WIP turns at the beginning of the improvement process, i.e., time "t"=0:

$$\text{WIP Turns Improvement} = \frac{\left(\dfrac{\text{Cost of Goods Sold}_0}{\text{Work In Process}_0}\right)}{\left(\dfrac{\text{Cost of Goods Sold}}{\text{Work In Process}}\right)} = \left(\frac{(W)(COGS_0)}{(W_0)(COGS)}\right)$$

[0053]   A Complexity Reduction Value Stream Map (CVSM) begins by using colors or other techniques to visually separate value-added from non-value-added work on a process map.   An example of a CVSM is shown in FIGURE 6.   Through a series of analyses, additional data summarized below can be added.

[0054]   A cross functional team collects data in a CVSM which indicates where the waste is and provides the data needed to eliminate the waste, improve quality and speed. The data will be collected separately at each activity within a process.   The average and the standard deviation of all data may be collected.

[0055]   Classification of activities as (a) value add, (b) non-value add, (c) required non-value add, (d) process time (P/T) (Value add time per unit for each type of service or product), and (e) time between changing from one service or product to another should be carried out.   Data should be collected on all activities in which item (d) is greater than item (e).   Still further activities to be classified are (f) changeover or setup time (time to switch between tasks), (g) queue time (the time "things" spend waiting to be processed),(h) take time (the demand rate of customers for each type of service or product), (i) complexity (number of different services or products processed at the activity), (j) uptime (time worked per day less breaks and interruptions), and (k) defects and rework percent per service or product at each activity with time and cost.

[0056]   From this activity level, data regarding current and future process performance can be computed including process cycle efficiency, lead time and variation, prioritized time traps, cost reduction and revenue growth opportunity.

[0057]   The starting value of non-value add cost, and the ending residual non-value add cost are estimated from the mapping process.   The Complexity Value Stream Map parses activities between value add and non-value add and is business enterprise specific.   Some portions of non-value add cost, such as supervision, planning, depreciation and

amortization, and stockroom costs may be reduced, but not eliminated whereas non-value add activities such as rework, scrap, setup etc may be eliminated to a greater degree. The result is that some of the non-value add costs are irreducible as a function of WIP. The amount of non-value add may be computed from a CVSM. If one is able to compute this impact on Work In Process separately, rational investments can be made in Six Sigma, Lean, and product value optimization and standardization efforts.

[0058] This invention also provides a method for computing and prioritizing the process drivers of Work In Process, among which are product complexity, quality problems, setup time, etc. By prioritizing these drivers of WIP, costs defined in a CVSM can be eliminated. The amount of WIP cannot be reduced by management fiat. Rather, specific process improvements are required. An example is that of a large press having a long setup time. To maintain productivity, each product will be produced in large batches prior to another setup. This necessarily creates large amounts of Work In Process. Editing low Work In Process would result in producing in batch size of one with catastrophic results to cost. Rather than management fiat, the process improvement known as the Four Step Rapid Setup method must be applied (e.g. Shingo). This invention derives a method for computing the intrinsic or minimum Work In Process that is currently required to meet existing

demand, and use this output to prioritize reductions in product complexity, setup time, downtime, quality problems etc.

[0059]   The model contained herein describes a general batch-flow manufacturing or service/transactional environment in which batches of distinct parts or work products are processed in serial fashion through distinct workstations or activities.   Using the analysis of prior U.S. Patents 5,195,041 and 5,351,195, which are incorporated herein by reference, and extending the analysis to the complexity of the product line, the minimum batch size consistent with meeting customer demand without delay or stockout, and the workstation turnover time implied by that batch-size regimen, may be derived.   The analysis is then extended to derive formulas for the average aggregate (i.e., across parts/products) Work in Process Inventory and the average aggregate Finished Goods Inventory that result from this processing strategy.   While the derivation is performed in a manufacturing context, it will be shown to apply to any process.

[0060]   The equations of the present invention can be used for the following purposes:

1.   To determine the size and location of Work In Process inventories in the process.

2.   To determine the effects of set-up times, downtimes, poor quality, and part (product) complexity

on the average aggregate inventories in the facility. This permits the user to focus efforts on, for example, setup time reduction, quality improvement, and part/product simplification at those points in the process in which the reduction in inventories will be maximized. It also allows the user to predict the magnitude of inventory reduction that will be achieved by a specified setup-time reduction, downtime prevention, quality improvement, or part/product complexity reduction, thus permitting the user to set priorities in improvement efforts.

3. To reduce the overhead cost, quality cost, and lost revenue that are related to large amounts of Work In Process and the resulting slow cycle times.

4. To construct a computer software application that permits these equations to be used in a plant of arbitrary size (i.e., number of workstations) and arbitrary complexity (i.e., number of distinct parts/products).

[0061] A product manufacturing example is a suitable example because the resulting work in process inventory can be physically observed and traced through the process. The conclusions apply with equal force to service process applications such as procurement, marketing, and product development, whose "Work In Process" now generally exists as

electronic files moving through a network of computer work stations and storage media.

[0062]   Consider a manufacturing facility that processes parts and subassemblies in batches through a series of sequential workstations, as shown in FIGURE 7. For ease of specification, the same processing route may be provided for all parts and subassemblies.   In accord with predominant manufacturing practice, the flow may be modeled as possessing the following characteristics:

1.) Raw materials and/or purchased parts are delivered to the upstream terminus of the process in batch quantities of the same size as the processing batches required at a first workstation.

2.) Batches remain integral throughout the process. This characterization implies two separable features. First, batches are not re-sized to accommodate different processing rates and/or setup times at each workstation.   Second,   as parts   complete their processing at an upstream workstation, they are transported to the next workstation in transfer batches of the same size as the processing batch.

3.) Parts are transported from the final workstation (J) to a finished goods inventory storage location, to await delivery to customers.   Parts/products are transported in transfer batches of size equal to the processing batch at final workstation J, FIGURE 7.

However, customers are not constrained to accept
delivery in batches or in production sequence. One can
model the demand for finished goods as simultaneous,
parallel flows for all parts/products, each with a
characteristic demand rate expressed in units of
product per unit time.

For example:

$$\Lambda = \sum_{i=1}^{N} \lambda_i \tag{1}$$

where:

i   =   part/product index, i = 1, …, N. That is, the
                facility produces N different parts or products.

λ_i =   customer demand rate for part/product i, expressed
                in units per hour. Thus, λ_i is an element in an N-
                dimensional vector of demand rates.

Λ =   aggregate customer demand for all parts/products
                produced in the facility.

                Certain ones of the parameters identified below
                are used in further equations which follow herein:

S_{ij} =   the setup time required to prepare workstation j
                to produce a batch of part/product i. Thus, S_{ij} is
                an element of an N x J matrix of setup times.

P_{ij} =   the time required to process one unit of
                part/product i at workstation j. Thus, P_{ij} is an
                element of an N x J matrix of processing times.

B_{ij} =   the batch size in which part/product i is
                processed at workstation j. Thus, B_{ij} is an
                element of an N x J matrix of possible batch
                sizes.

j   =   workstation index, j = 1, …, J. That is, the
                processing route consists of J distinct
                workstations.

[0063]   Using the logic and equations developed in the
methods described in U.S. patents 5,195,041 and 5,351,195,

21

one can derive the minimum batch size that can be processed
at each workstation that is consistent with meeting the
vector of demand rates without stockout or delay. This
minimum batch size may be denoted by $MINB_{ij}$. In other words:

$MINB_{ij}$ = minimum batch size allowable for processing
part/product i at workstation j, which fulfills
demand rate λ_i.

[0064] Processing of the minimum batch sizes $MINB_{ij}$ at
workstation j implies a workstation turnover time (WTT) for
that station, i.e., the time interval required to complete
one production cycle through all N products. This time
interval may be denoted by WTT_j, wherein:

WTT_j = workstation turnover time at workstation j, within
which the workstation can be setup for, and
complete processing of one batch of size $MINB_{ij}$ for
all N parts/products produced by the facility.

[0065] The algorithm described in U.S. patents 5,195,041
and 5,351,195 provides a method to compute MINBij and WTTj
for each product i and each workstation j. That is, the
algorithm and an associated computer program may compute the
minimum batch size consistent with meeting final customer
demand without stocking out of any part or product, as well
as the workstation turnover time associated with the minimum
batch sizes. However, this procedure does not compute the
implied work in process and finished goods inventory levels
implied by this strategy. That extension is the focus of
the methods described herein. Given the description of the
processing system provided, and utilizing the minimum batch

size computations, it is possible to derive the Work In Process and finished goods inventory profiles.

[0066]   The amounts of Work In Process and finished goods inventories are determined both by the physical parameters of the process (e.g., processing times and setup times) and by management policies, such as the production schedule. In the derivations that follow, the inventory levels may be derived that are induced by (a) using the minimum batch sizes consistent with meeting customer demands, and (b) considering a range of production schedules that vary from perfectly synchronized production to fully asynchronized production; these terms are defined below.

PERFECTLY SYNCHRONIZED PRODUCTION

[0067]   To illustrate the inventory profiles, consider the following production schedule. In the table of FIGURE 8, the first two columns are the time index, and the three further columns depict the status of three serial workstations. The label SA indicates that the workstation is being setup to process part A, and the label PA indicates that the workstation is processing a batch of part A. Identifiers SB, PB, SC and PC indicate the corresponding activities for parts B and C, respectively. One complete workstation turnover cycle includes setting up for part A, processing a batch of part A, setting up for part B, processing a batch of part B, setting up for part C, and processing a batch of part C.

[0068] This schedule exhibits perfectly synchronized production. While this condition is ideal, and seldom if ever realized, it provides a transparent derivation needed for the asynchronous processes which are observed in practice. That is, the batch release timing has been coordinated so that, for example, Workstation 2 of FIGURE 7 has been setup to accept a completed batch of part A from Workstation 1 just as Workstation 1 completes that batch of part A. This same "smooth handoff" occurs for all products at all workstations. Similarly, raw material deliveries occur in perfect just-in-time fashion, i.e., immediately following the setup procedures required to prepare Workstation 1 to process the part in question.

[0069] In general, setting the minimum batch size of all stations to that implied by the most constrained workstation will facilitate this smooth flow, although we should emphasize that perfect synchronization cannot be guaranteed in all situations. It is possible to construct examples in which the particular sequences of setup times and processing times do not permit perfect synchronization without occurrences of starvation or blocking between at least one pair of consecutive workstations, hence emphasizing the need for a detailed computational program. However, the combination of minimum batch sizes with a perfectly synchronized schedule minimizes work in process inventories, thus providing a benchmark for comparative purposes.

[0070] This simplified scenario can be continued to depict the inventory buildup and usage profiles at typical workstations in the facility. To ease interpretation of FIGURES 9 through 12, the timing labels shown in FIGURE 8 have been maintained and the minimum batch sizes of parts B and C have been set to be twice and three times that of part A, respectively.

[0071] Consider the first (i.e., furthest upstream) workstation in the routing, labeled WS1. Raw materials or purchased parts are delivered to WS1 just as they are needed to begin processing, i.e., just as the appropriate setup operation has been completed. These raw materials are converted to Work In Process (WIP) at WS1, hence the buildup of WIP as raw material inventory declines. The now partially processed WIP is transferred to Workstation 2, and the pattern of incoming WIP decline and outgoing WIP buildup continues, as the batches flow downstream through the series of workstations. This policy yields the inventory profiles shown in Figures 9 and 10.

[0072] After passing through the first J-1 workstations, the batches enter the final workstation before finished goods inventory. The inventory profile changes here, because a batch size sufficient to meet all customer demand throughout the workstation turnover time interval must be processed. This batch is processed at workstation J, and enters finished goods inventory, from which it is depleted

25

as deliveries are made to customers at rate λi for each of the N products. These deliveries are made in parallel (i.e., simultaneously) for all products. This policy leads to the inventory profile shown in Figure 11.

[0073] The algorithms described herein can be used to compute the inventory levels at the various workstations in the facility, and to aggregate those values to compute the aggregate inventory levels maintained in the facility. Figure 12 shows aggregate inventory levels for a simplified production example of synchronized production.

[0074] Real factories that have been set up with "synchronous manufacturing" in mind have in fact minimized work in process inventories and have done so at the expense of large finished goods inventory. In attempts to apply "synchronous manufacturing" using a variant, without improving setup time, etc. it has been observed that automotive engine factories in which the whole factory was setup to machine and assemble engine type A that Work In Process flows from index machining to assembly and test was carried out with minimum Work In Process. In one case, four engine types were produced, with three days spent per engine type. The Cycle Time Interval between repetitions was twelve days. Assuming demand exactly matches supply, an average of at least six days of inventory exists in finished goods. Thus the non-value add cost has in part been exported to the cost of a warehouse.

[0075] As noted above, the perfectly synchronized production schedule represents an ideal case that may not be attainable, even in theory, for a given set of production parameters (setup times and processing times). The "next best" scenario is shown in the table of FIGURE 13. In this schedule, the setup procedures for part A at Workstation 2 (WS2) begin when the batch of part A arrives at Workstation 2 for processing, and similarly at all workstations, including Workstation 1 (WS1). The schedule actually assumes that setup for Part A at Workstation 2 begins as the batch of part A is completed at Workstation 1, acknowledging that the physical transport of the batch between stations can constitute part of the setup of the downstream station.

[0076] The setup-on-batch-arrival scheduling policy increases the amount of Work In Process and raw material inventories in the facility, as shown in Figures 14 and 15. The amount of finished goods inventory is unaffected, although its timing is shifted, as shown in Figure 16.

[0077] Figures 14 and 15 indicate that the raw material and Work In Process inventory profiles are trapezoids rather than triangles. That is, batches do experience periods of waiting in which no units in the batch are being processed. Still, this schedule is highly synchronized, and the amount of waiting is minimal. Although somewhat less likely, it also is still possible to construct a hypothetical set of products, setup times, and processing times for which no

setup-on-batch-arrival schedule can be constructed without occurrences of starvation or blocking between at least one pair of consecutive workstations.

[0078] The profile of total inventory in the facility is shown in Figure 17. The shape of the profile is identical to that of Figure 12 (i.e., for perfect synchronization), but the absolute levels are increased. That is, the level of total facility inventory varies between 18.5 and 27.5 units under a perfectly synchronized schedule, but the Setup-On-Batch-Arrival Schedule yields a total inventory level that varies between 23 and 32 units.

MODERATELY ASYNCHRONIZED PRODUCTION

[0079] To depict a schedule that departs even further from the perfectly synchronized example above, and represents a real process, consider the schedule shown in FIGURE 18. In this schedule, the various batches are out of perfect synchronization by one part/product. That is, when a batch of part A arrives at Workstation 2, it must wait for a batch of part C to be setup and processed before it can be setup and processed. While clearly sub-optimal from an inventory minimization perspective, this schedule might be viewed as the upper limit of observed best practice in a manufacturing or transaction processing environment.

[0080] The raw material and work in process inventory profiles in Figures 19 and 20 continue to be trapezoidal in shape, but now the waiting periods between processing steps

have grown, as indicated by the extended bases of the trapezoids. Inventory management itself now becomes more complex. Consider the time interval from hour 6 to hour 7 in Figure 20. During this period, batches of all three products are present between Workstations 1 and 2; a batch of part C still is being processed at Workstation 2, a batch of part A is awaiting processing at Workstation 2, and a batch of part B is building up at the downstream port of Workstation 1.

[0081] The profile of total inventory in the facility is shown in Figure 22. Again, the shape of the profile is identical to those shown in Figures 12 and 17, but the absolute levels have increased further, now varying between 41 and 50 units.

[0082] As the figures associated with the preceding two forms of production suggest, raw material and Work In Process inventories continue to grow as the production schedule departs further from the ideal of perfect synchronization. To provide a theoretical benchmark at the other end of the spectrum, consider the schedule shown in FIGURE 23. Under this scheduling regime, all workstations are performing the setup associated with product A at the same time; similarly for products B and C. Thus, when a batch of product A arrives for processing at Workstation 2, it finds itself last in line behind one batch of every other product. If one were to allow the schedule to deteriorate

further, a watchful scheduler would notice that each upstream workstation could be shut down for an entire workstation turnover cycle, to allow one set of batches of all products to "flush through" the system, to reduce inventory levels without degrading delivery service to customers. Thus, this schedule suggests the maximum level of inventory that would be encountered before alternative corrective actions would be anticipated.

[0083] The shapes of the raw material and Work In Process inventory profiles are similar to those of the preceding section, with the bases of the trapezoidal sections further enlarged. As before, the shape of the aggregate inventory profile is unchanged, but its level has grown, now varying between 59 and 68 units, roughly three times that of the perfectly synchronized case, as shown in FIGURE 24. Real processes typically operate between Case 3 and Case 4, shown by comparison in FIGURE 25.

[0084] Using the variables defined for equation (1), it is possible to derive equations to determine the Work In Process and finished goods inventory levels in the facility. One may begin by determining the maximum of the workstation turnover times required to produce one (minimum-size) batch of each product at each workstation.

$$ WTT_{max} = \underset{j=1,J}{Max} \left[\frac{\sum_{i=1}^{N} S_{ij}}{1 - \sum_{i=1}^{N} \lambda_i P_{ij}} \right] \qquad (2) $$

[0085] Under the generally applied policy of maintaining a consistent processing and transfer batch size for each part/product as it traverses its processing route, one can determine the batch sizes for the N distinct parts/products:

$$\text{MINB}_i = \lambda_i \text{WTT}_{max} = \lambda_i \max_{j=1,J} \left[\frac{\sum_{i=1}^{N} S_{ij}}{1 - \sum_{i=1}^{N} \lambda_i P_{ij}} \right] \tag{3}$$

[0086] The inventory profile diagrams and the processing schedules can be used to determine the implied work in process inventory at each workstation, and sum over the workstations to determine the total work in process inventory of part/product i in the facility. These are the inner summations over the index j in equations 4, 5, and 6. Note that one must carry forward the summations in equation (3), embodied in the term MINB_i. Summing over the part/product index i yields the average aggregate Work In Process (WIP) inventory in the facility.

Perfectly Synchronized Average WIP =

$$\sum_{i=1}^{N} \left[\lambda_i \sum_{j=1}^{J} \left(\text{MINB}_i P_{ij} \right) \right] \tag{4}$$

Setup-On-Batch-Arrival Average WIP =

$$\sum_{i=1}^{N} \left[\lambda_i \sum_{j=1}^{J} \left[S_{ij} + \left(\text{MINB}_i P_{ij} \right) \right] \right] \tag{5}$$

Fully Asynchronized Average WIP =

$$\sum_{i=1}^{N}\left[\lambda_i\sum_{j=1}^{J}\left(\sum_{i=1}^{N}\left[S_{ij}+\left(MINB_iP_{ij}\right)\right]\right)\right] \qquad (6)$$

[0087]   In a similar fashion, one can use the finished goods inventory profiles and the processing schedules can be used to derive the average aggregate level of finished goods on hand.   As indicated in the discussion of the finished goods inventory profiles, the average finished goods inventory is unaffected by the level of synchronization of the production schedule.

Average Finished Goods Inventory =

$$\sum_{i=1}^{N}\frac{1}{2}\lambda_i WTT_{max} \;=\; \sum_{i=1}^{N}\frac{1}{2}MINB_i \qquad (7)$$

[0088]   Combining equations 4, 5, and 6 with equation 7 yields the average total inventory in the facility under each of the scheduling scenarios.   Again, the value of $MINB_i$ must be computed from equation 3 and inserted in equations 8, 9, and 10.   Note that these equations could be factored into a slightly more compact notation, but this would conceal their basic structure as examples of Little's Law, and reduce the comparability across scheduling scenarios.

Perfectly Synchronized Average Total Inventory in System =

$$\sum_{i=1}^{N}\left[\lambda_i\sum_{j=1}^{J}\left(MINB_iP_{ij}\right)\right] + \sum_{i=1}^{N}\frac{1}{2}MINB_i \qquad (8)$$

Setup-On-Batch-Arrival Average Total Inventory in System =

$$\sum_{i=1}^{N}\left[\lambda_i\sum_{j=1}^{J}\left[S_{ij}+\left(MINB_iP_{ij}\right)\right]\right] + \sum_{i=1}^{N}\frac{1}{2}MINB_i \qquad (9)$$

Fully Asynchronized Average Total Inventory in System =

$$\sum_{i=1}^{N}\left[\lambda_i \sum_{j=1}^{J}\left(\sum_{i=1}^{N}\left[S_{ij}+\left(MINB_i P_{ij}\right)\right]\right)\right] + \sum_{i=1}^{N}\frac{1}{2}MINB_i \qquad (10)$$

[0089] The complexity of the multiple summations may mask the effects of set-up time and product complexity on the amount of work in process and finished goods inventory maintained in the facility. The roles of these parameters can be highlighted by considering a system in which all parts/products and all workstations are symmetric, i.e., in which the set-up times and processing times per unit are identical for all products and all workstations.

$$\lambda_i = \lambda \quad \text{for all i} \qquad (11)$$

$$P_{ij} = P \quad \text{for all i, j} \qquad (12)$$

$$S_{ij} = S \quad \text{for all i, j} \qquad (13)$$

[0090] Recall that the aggregate demand (across all products) is denoted by the symbol Λ, which was defined in equation (1). With these simplifications,

$$\Lambda = \sum_{i=1}^{N}\lambda_i = N\lambda \qquad (14)$$

[0091] Also note that in order to deliver Λ units of product per hour, each of which requires P hours of processing at each sequentially ordered workstation, one must have:

$$0 \leq \Lambda P \leq 1 \qquad (15)$$

33

[0092] These parameter values can be substituted to determine the common minimum batch size for all products and the common workstation turnover time for all workstations.

$$MINB = \lambda \left[\frac{\sum_{i=1}^{N} S}{1 - \sum_{i=1}^{N} \lambda P} \right] = \frac{\Lambda S}{1 - \Lambda P} \qquad (16)$$

$$WTT = \left[\frac{\sum_{i=1}^{N} S}{1 - \sum_{i=1}^{N} \lambda P} \right] = \frac{NS}{1 - \Lambda P} \qquad (17)$$

[0093] By successive simplifications(in which all products have the same demand, setup time, processing time/unit, etc.), one can demonstrate the effects of setup time S and part/product count N. Given the definitions:

Scheduling Case= [Work In Process]+ [Finished Goods]

Perfectly Synchronized Average Total Inventory in System =

$$\left[\frac{(\Lambda P)J\Lambda S}{1 - \Lambda P} \right] + \left[\frac{N\Lambda S}{2(1 - \Lambda P)} \right] \qquad (18)$$

Setup-On-Batch-Arrival Average Total Inventory in System =

$$\left[\frac{J\Lambda S}{1 - \Lambda P} \right] + \left[\frac{N\Lambda S}{2(1 - \Lambda P)} \right] \qquad (19)$$

Fully Asynchronized Average Total Inventory in System =

$$\left[\frac{NJ\Lambda S}{1 - \Lambda P} \right] + \left[\frac{N\Lambda S}{2(1 - \Lambda P)} \right] \qquad (20)$$

[0094] From these equations, one can see that the average total inventory in the facility is linear in the setup time S, under all scheduling scenarios. Moreover, the setup time affects the level of both work in process inventories and

finished goods inventories. Thus application of the Four Step Rapid Setup Method (Shingo) to radically reduce setup time will reduce inventory and the related non-value add costs. Because setup time is a common factor, its reduction ameliorates the impact of N, J,P and other parameters. However, the effort involved in effecting radical setup reduction when N and J is large can·be prohibitive.

[0095] The average Work In Process inventory is linear in the number of distinct workstations, J. Each workstation represents both a position at which a batch of parts/products should be undergoing processing at all times, and a point at which batches awaiting processing can accumulate. One of the fundamental tools of the Lean process improvement initiative is Value Stream Mapping which has been generalized to Complexity Reduction Value Stream Mapping herein. This tool allows improvement activities to then be focused on the non-value add activities and effect their elimination, thus reducing J. As discussed below, real processes tend toward equation(20),hence the reduction of J can significantly reduce inventory and related non-value add costs as well as focusing setup reduction on the reduced number of activities.

[0096] The complexity of the facility's part/product portfolio plays a somewhat more subtle role in the equations. The overall total demand rate faced by the facility is represented by the symbol Λ:

$$\Lambda = \sum_{i=1}^{N} \lambda_i = N\lambda \qquad (21)$$

[0097]   One form of complexity can be modeled as the number of distinct or differentiated parts/products into which the overall demand rate λ has been segmented. Here the operational definition of differentiation is a feature of a part or product that necessitates a setup operation at each workstation before a batch can be processed. Thus the magnitude of N denotes the processing complexity. As the right-hand term in equations (18) through (20) indicates, the average finished goods inventory level in the facility is linear in N. Finished goods inventory must be carried in each part/product, and each distinct product contributes another setup and processing interval to the workstation turnover time. Sufficient finished goods must be stocked to satisfy demand over that workstation turnover time.

[0098]   Equations (18) and (19) indicate that, if the processing schedule is perfectly synchronized or at least very close to perfectly synchronized (setup-on-batch-arrival), the Work In Process inventory is independent of N. The only known case of near perfect synchronization occurred in the Highland Park Model T plant in which N=1. However, as N grows, it becomes more difficult to devise and maintain a perfectly synchronized schedule, and thus the Work In Process inventory will grow. Equation (20) more accurately represents reality, and indicates that the

benchmark value of WIP inventory for a fully asynchronized schedule is linear in N in both WIP and Finished Goods, which can be very much greater. Therefore complexity reduction has a major impact on WIP and hence non-value add cost reduction.

[0099]   Increased Complexity N creates large inventory though its impact on P and S. As N increases, the period between repetitions of any given product increases. If a product is built every hour, learning does not decay. If it is built once per month, a relearning process may result. Thus, as N increases, both the setup time and the processing time per unit increases. The latter can be quite explosive, as it enters into the denominator as $1-\Lambda P$ which approaches zero as ΛP approaches 1 and hence WIP and Finished Goods approach infinity.

[0100]   To reduce finished goods inventory and reduce the possibility that Work In Process inventory is unnecessarily large, one could pursue a program of complexity reduction in two ways. First, one can standardize the product portfolio by redesigning products and processes to maximize part commonality across products.   One might model such an initiative as reducing N without reducing λ. That is, the total demand rate faced by the firm would continue at rate λ units of product per hour, but with fewer different parts embodied in those products.   In that case, one might expect

37

a linear reduction in finished goods inventory and a simplified scheduling regime.

[0101] Second, one could rationalize the product portfolio by ceasing to supply those products whose individual demand rates λi are small, but whose processing ultimately requires increased inventory levels of all products (not just product i) to provide sufficient productive capacity for their manufacture. In this case, equations (18) through (20) still are indicative of the inventory reductions that could be achieved, but it can no longer be assumed that the aggregate demand rate λ would remain constant.

[0102] Equations (18) through (20) made the simplifying assumption that all products had the same demand, setup time, etc. But what is the impact of Pareto Demand, in which 80% of the different products or services only constitute 20% of the demand? To conduct an analysis of a typical manufacturing facility or service process, one must use the specific setup time, processing time, and demand which differs for each product or offering. Due to the complex nature of the offering, this may run to thousands or tens of thousands of pieces of data, most of which are available on ERP systems. A computer program may be constructed to execute the logic embodied in equations (2) through (6) which allow real demand, setup time, etc to be used. Such a program, which may be called Supply Chain Accelerator (SCA),

38

will calculate the minimum safe batch size (MBS), and the average total WIP. Inherent in equations (2) through (6) is the assumption of perfect synchronization of the facility, and the examples given herein focus on a small number of similar parts, allowing this perfect synchronization to be achieved. However, in the case of a more complex facility with many parts and workstations, all with very different demand rates, processing and setup rates, scrap rate, etc., the assumption of perfect synchronization of the facility can no longer be assumed. In this case, the deterministic model of the facility outlines in equations (2) through (6) must give way to a stochastic model that is able to capture the imperfections in scheduling that occur in a complex factory.

[0103] To build these stochastic models of facilities, a second software program called Process Cycle Simulator (PCS) may be developed. PCS relies on discrete event simulation to simulate the running of the facility for many years. Statistics are collected indicating the level of WIP and finished goods that are held, as well as the service level attained. The simulation models are started with the minimum safe batch sizes determined by SCA, with a portion of capacity held in reserve in the SCA model to allow for the imperfections of scheduling in a complex facility. PCS is run with the output of SCA, the statistical results are observed, and if necessary, additional runs of the PCS are

executed with modified minimum safe batch sizes in order to arrive at the proper balance of service level and WIP desired by management.

[0104]   PCS also has the ability to add random variation to all of the input parameters of equations (2) through (6). This allows the user to add another dimension of reality to the modeling exercise, namely variation. As all processes experience some degree of variation, the ability of PCS to model this variation and determine its effect on service level and WIP further strengthens the suite of tools developed to model, understand, and attack complexity.

[0105]   To demonstrate this process, a PCS model may be built of a facility that represents only a small increase in complexity over the facility that was the basis of equations (2) through (6). In this facility, there is still only a single workstation, but the number of parts being processed by that workstation have been increased to sixteen. Furthermore, while the setup and processing times remain constant for all parts in the facility, two different demand rates are utilized. The low demand parts – "Slow Movers" – are each given just 1.5% of total demand. The number of Slow Movers is varied from one to fifteen. The high demand parts – "Fast Movers" – are given the remainder of the demand to be split equally amongst the parts. The range of models considered was the following: two to fifteen Fast Movers, and zero to fourteen Slow Movers. Each

combination of these parameters was examined, as shown in FIGURES 26 and 27. From the table of FIGURES 26 and 27, the dramatic effect of the addition of just this small amount of complexity can be seen. For example, in the case of five Fast Movers and no Slow Movers, the base WIP calculation is 2500. Recall that demand is kept constant across all parts. So, with total demand unchanged, the addition of eleven Slow Mover parts to the factory causes the new WIP calculation to more than double to 5747. Thus these eleven low demand(slow moving) part numbers effectively double the Work In Process.

[0106] Referring to Figure 5, such an increase in WIP could drive the enterprise far up the Cost of Goods Sold curve. This in turn would mean that the related costs as determined in a CVSM would have to be borne by the eleven products. Activity Based Costing would, in principle, come to the same conclusion, and might suggest that these eleven products were in fact not earning their cost of capital and should be eliminated. The point of departure is based on the strategy of continuous improvement. Cost of complexity reduction would suggest that the eleven products be evaluated in terms of potential value creation. Assume that all setup times could be cut in half. According to Equation (20) WIP and Finished Goods would be cut in half, moving the Non-value add Cost of Goods Sold down the curve. If at this lower overhead rate a product could more than earn its cost of capital, it should be retained. Otherwise it is a

candidate for outsourcing or removal. Thus a focused process improvement process can yield higher value when directed by cost of complexity analysis that is hidden from both traditional accounting and activity based costing.

[0107]   Recall that the values in FIGURE 8 represent the WIP generated under the assumption of perfect synchronization of schedules. Figure 28 shows the response surface generated by the values in FIGURE 8.

[0108]   Using discrete event simulation without Process Cycle Simulator, the effects of synchronization imperfects and variation can be observed. For the case of three Fast-Mover parts and two Slow Mover parts, the base calculation of WIP with perfect synchronization and no variation is 1903 pieces. In this case, the backlog of demand that was not able to be fulfilled is zero, and the service level is 100%. Figures 29 and 30 show the effect on average backlog (average unfulfilled orders) when synchronization and variation are added to the model.

[0109]   Figure 29 clearly shows that even with no variation (Demand COV or Coefficient of Variation = 0), an average backlog of about seventy five pieces results purely from synchronization difficulties. When demand variation is added as seen by the Demand COV values on the x-axis, the average backlog increases dramatically. Since these simulations were performed with the minimum safe batch sizes calculated from equations (2) through (6), the WIP value has

remained relatively constant (with a small allowance for the variation in the model) at 1903, yet the service level has dropped dramatically. In order to return to a service level of approximately 98%, one would need to add WIP equal to three times the average backlog amount. This is based on the assumption of normality in the variation of demand, and three standard deviations to the right of the mean encompassing approximately 98% of the area of the normal curve. In the case of Figure 29 with a COV of 0.5, that would indicate raising average WIP from 1903 to about 2500 pieces.

[0110] Figure 30 shows the effect of setup and processing time variation on average backlog. Here, the effects of variation are even more pronounced that those observed for demand variation. Looking again at the case of the COV = 0.5, the average backlog becomes about 1200 pieces. Relying on the assumption of normality once again, the average WIP would need to increase from 1903 to about 5500 pieces to maintain an approximate 98% service level guarantee.

[0111] The facility modeled by these simulation exercises represents a very small increment in complexity, three Fast Mover parts and two Slow Mover parts, relative to situations that are typically encountered in industry where one may easily be working with hundreds or thousands of parts across dozens of workstations. Clearly, the adverse effects of complexity are visible in equations 2 through 6. When the

assumption of perfect synchronization is relaxed, the effects become even worse. When the reality of variation is then added to the modeling exercise, the effects of complexity become nearly intolerable. Yet, many businesses face this challenge on a daily basis. The purpose of the equations derived herein, are to help determine the magnitude of the problem that a facility faces, determine the best course of action to correct that problem, and estimate the benefits that the facility will enjoy once the complexity problem is resolved.

[0112] The equations have been developed using easy-to - visualize manufacturing examples. But the methodology and Little's Law apply to all processes. Consider the application to procurement. A centralized procurement operation purchased products for fourteen separate operating units each with its own legacy computer system. Each operating unit called a given component by a different name, used different(multiple) screens from which to buy, etc. A commodity buyer would log on to a given units computer, then re-learn the process of procurement. During this learning process, productivity would be low. Because this learning period acted like setup time, the buyer would stay locked on to an operating units requirements until all were purchased, typically taking up to two days. With fourteen operating units, the Cycle Time Interval was twenty-eight days. Ideally the buyer would only place the required "hot"

orders, and then log on to another site. This would have led to very low productivity and the need to hire more buyers. Low productivity would mean a longer Processing time per unit P. Figure 31 depicts the process. According to equation (20) this non-value add cost would have been disclosed by larger inventory since $1-\Lambda P$ appears in the denominator. The solution chosen was three fold:

1. Eliminate Setup time: a database was setup to poll all operating units during the evening hours and prepare a corporate wide "hot list" for the buyer, with consolidated buys leading to 10% lower material cost and the number of buyers reduced 50%.

2. Standardization: The process of buying was made uniform by placing commodities on "contract" such that the supplier had already agreed to price and delivery, making the buy a "point and click" operation rather than a different task for each buy.

3. Cycle time: by only buying "hot" items, the Work In Process by a factor of twenty and cycle time to serve each operating unit was reduced below two days.

[0113]   The complexity was not reduced. The same number of different part numbers existed. But the transparent, non-value add, cost of buying each commodity separately was

eliminated, and the cost of each commodity was reduced by standardization, reducing the cost of complexity.